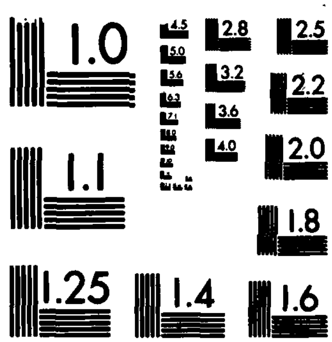
AD-A160 980 IMPLICIT B-DIFFERENTIABILITY IN GENERALIZED EQUATIONS (U) WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER S M ROBINSON AUG 85 MRC-TSR-2854 DAAG29-80-C-0041 1/1

UNCLASSIFIED NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

MRC Technical Summary Report #2854

IMPLICIT B-DIFFERENTIABILITY IN GENERALIZED EQUATIONS

Stephen M. Robinson

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53705

August 1985

(Received July 8, 1985)

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, DC 20550

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

IMPLICIT B-DIFFERENTIABILITY IN GENERALIZED EQUATIONS

Stephen M. Robinson

Technical Summary Report #2854
August 1985

ABSTRACT

~~We have~~ previously proved an implicit-function theorem for regular solutions of generalized equations. Here ~~we~~ show that when the underlying set for the generalized equation is polyhedral, as it is in many applications, then the implicit function has a Bouligand derivative defined by a formula generalizing that of the usual implicit-function theorem. This extends recent results on directional differentiability obtained by Kyparisis and others.

AMS (MOS) Subject Classifications: 90C33, 90C30

Key Words: Equilibrium, generalized equations, implicit functions, Bouligand derivative

Work Unit Number 5 - Optimization and Large Scale Systems

Sponsored by the National Science Foundation under Grant No. DCR-8502202, and by the United States Army under Contract No. DAAG29-80-C-0041.

-B-

SIGNIFICANCE AND EXPLANATION

Generalized equations are mathematical models for nonlinear equilibrium problems in areas such as economics, transportation, etc. In such models, it is desirable to know how the solution of the model will change when the problem data change. When such a change is too hard to compute, a convenient approximation method may yield an answer that is good enough, particularly for small changes. This paper develops such an approximation. See p -A-

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

IMPLICIT B-DIFFERENTIABILITY IN GENERALIZED EQUATIONS

Stephen M. Robinson

1. Introduction. This paper deals with solutions of generalized equations of the form

$$0 \in f(p,x) + \partial\psi_K(x) \quad , \qquad (1.1)(p)$$

where $f : \Pi \times \Omega \to \mathbb{R}^n$, Π and Ω are open subsets of a normed linear space P and of $\mathbb{R}^n$ respectively, and K is a polyhedral convex set in $\mathbb{R}^n$. The variable p is a perturbation parameter; for fixed p, (1.1)(p) expresses the geometric requirement that $f(p,x)$ be an inward normal to K at x. Generalized equations like (1.1)(p) can be used to model a wide variety of equilibrium and optimization problems; for discussion and examples, see the survey in [3].

In [2, Th. 2.1] we showed that if for a fixed $p_0 \in \Pi$ a solution $x_0 \in \Omega$ of (1.1)(p) possesses a certain property that we shall call regularity, then for p near p_0 (1.1)(p) defines a (locally) single-valued implicit function $x(p)$ having the property that the pair $(p,x(p))$ solves (1.1)(p). Further, in [2, Cor. 2.2, Th. 2.3] we showed that under mild additional assumptions $x(\cdot)$ is Lipschitzian and is closely approximated by the solution $z(p)$ of the linear generalized equation

$$0 \in f(p,x_0) + f_x(p_0,x_0)(z-x_0) + \partial\psi_K(z) \quad , \qquad (1.2)$$

where $f_x(p_0,x_0)$ denotes the partial Fréchet derivative with respect to x. All this is reminiscent of the implicit-function theorem, yet there is one aspect of that theorem that does not appear in [2]: the fact that if the original function is Fréchet differentiable then so is the implicit function, and that its derivative may be calculated by implicit differentiation. The reason this result does not appear in [2] is that it is not true for generalized equations: one may show easily by examples that even if the solution of (1.1)(p) is regular it may not be Fréchet differentiable at p_0.

More recently, in a paper primarily concerned with variational problems over perturbed sets, Kyparisis used results of [2] to prove that if f is C^1 at (p_0,x_0) then the

---

Sponsored by the National Science Foundation under Grant No. DCR-8502202, and by the United States Army under Contract No. DAAG29-80-C-0041.

solution $x_D(p)$ of

$$0 \in f(p_0,x_0) + f'(p_0,x_0) \begin{bmatrix} p & -p_0 \\ x_D(p)-x_0 \end{bmatrix} + \partial\psi_K(x) \tag{1.3}$$

approximates $x(p)$ to better than linear order in $\|p-p_0\|$ [1, Lemma 4.1]. He actually proved this for the case in which $\Pi \subset R^k$ and K is a closed convex (not necessarily polyhedral) set. He then observed that this result implies directional differentiability of x at p_0 if x_D is directionally differentiable there [1, Cor. 4.2], and he related this fact to several other papers on variational inequalities and equilibrium problems; see [1] for references to these papers.

Our aim in this paper is twofold. First, we show that when K is polyhedral $x(\cdot)$ has a stronger property than directional differentiability: in fact, it has a Bouligand derivative (B-derivative) at p_0. The B-derivative was introduced in [5]; it has properties weaker than those of the Fréchet derivative but stronger than those of directional derivatives. In particular, B-differentiability implies directional differentiability in all directions, but it also describes the relationships of the directional derivatives to each other. Our second aim is to show that this stronger result holds (again, for polyhedral K) under weaker hypotheses than those used in [1] for general K.

In the next section we review B-differentiability and show how under some circumstances a B-derivative can be represented by a sum of partial derivatives. Then in Section 3 we prove the implicit-function theorem that is our main result.

2. Review of B-differentiability. Here we review some facts about B-differentiability that we shall need in Section 3. For proofs, see the appendix in [5].

Given a Lipschitzian function g from an open subset Ω of $\mathbb{R}^n$ to $\mathbb{R}^m$, we say that g has a B-derivative at $x_0 \in \Omega$ if there is a (single-valued) function $Dg(x_0)$ from $\mathbb{R}^n$ to $\mathbb{R}^m$ such that (1) the graph of $Dg(x_0)$ is a cone, and (2) one has for x near x_0,

$$g(x) = g(x_0) + Dg(x_0)(x-x_0) + o(x-x_0) \quad . \tag{2.1}$$

Of course, if the graph of $Dg(x_0)$ is a subspace then we have the usual notion of Fréchet (F-) differentiability, so any F-differentiable function is B-differentiable, although the converse is of course not true.

In [5] we showed that if (2.1) held at all then it held for just one function $Dg(x_0)$: that is, the B-derivative is unique when it exists. We also showed that $Dg(x_0)$ inherits the Lipschitz modulus of g, that it obeys the usual chain rule, and that B-differentiation is a linear operation in the space of Lipschitzian functions.

For our work in this paper we need to relate B-differentiability to partial differentiation. Of course one can define a partial B-derivative of a function of several variables: for example, in the case of $g(x,y)$ we would define $D_y g(x,y)$ to be the B-derivative of the function of y given by $g(x,\cdot)$. However, it should be clear that the usual formula representing the F-derivative of a function of several variables in terms of the individual partial derivatives will not work for B-derivatives. For example, consider the Euclidean norm $\|\cdot\|$ (which is its own B-derivative). Except for one-dimensional cases, this cannot be represented by sums of the partial B-derivatives of the components.

However, something can be salvaged if some of the components have continuous F-derivatives instead of just B-derivatives, as the following proposition shows.

PROPOSITION 2.1: Let g be a Lipschitzian function from $S \times T$ to $\mathbb{R}^k$, where S and T are open sets in $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively. Let $(x_0,y_0) \in S \times T$, and suppose that the partial Fréchet derivative g_x of g with respect to x is continuous at (x_0,y_0). Then

$$Dg(x_0,y_0)(v,w) = g_x(x_0,y_0)v + D_y g(x_0,y_0)(w) \quad . \tag{2.2}$$

Proof: By uniqueness of the B-derivative we need only show that the function on the right side of (2.2) has a cone for its graph, and that it approximates g to better than first order. The first statement is obvious. For the second, choose $\varepsilon > 0$ and let V (convex) and W be neighborhoods of x_0 and y_0 respectively, so small that if $x \in V$ and $y \in W$ then

$$\| g(x_0,y) - g(x_0,y_0) - D_y g(x_0,y_0)(y-y_0)\| \leq \varepsilon \|y-y_0\| \quad ,$$

and for each $\lambda \in [0,1]$,

$$\| g_x[(1-\lambda)x_0 + \lambda x,\, y] - g_x(x_0,y)\| \leq \varepsilon \quad .$$

Then

$$\begin{aligned}
&\| g(x,y) - g(x_0,y_0) - g_x(x_0,y_0)(x-x_0) - D_y g(x_0,y_0)(y-y_0)\| \\
&\quad \leq \| g(x_0,y) - g(x_0,y_0) - D_y g(x_0,y_0)(y-y_0)\| \\
&\qquad + \| g(x,y) - g(x_0,y) - g_x(x_0,y_0)(x-x_0)\| \\
&\quad \leq \varepsilon \|y-y_0\| + \| \int_0^1 \{g_x[(1-\lambda)x_0 + \lambda x, y] - g_x(x_0,y_0)\}(x-x_0)d\lambda \| \\
&\quad \leq \varepsilon(\|y-y_0\| + \|x-x_0\|) \quad .
\end{aligned}$$

Since ε was arbitrary, (2.2) follows immediately, completing the proof.

In applying Proposition 2.1, a change of coordinates may sometimes be useful. For example, consider the function $g(x,y) = x+y + \|x-y\|$. It is not hard to see that g does not have a partial derivative at the origin with respect to either x or y. However, by introducing the change of variable given by $w = x+y$ and $v = x-y$ one obtains the new function $h(v,w) = g[\frac{1}{2}(w+v), \frac{1}{2}(w-v)] = w + \|v\|$, to which Proposition 2.1 can be applied.

We shall apply Proposition 2.1 in the next section to separate the differentiability requirement on f with respect to x from those with respect to p. We shall see that although we require C^1 behavior in x, it suffices for f to have a B-derivative with respect to the parameter p. This will enlarge the class of functions to which our main theorem can be applied.

3. An implicit-function theorem. In this section we review the connection between regularity and the linearization of (1.1)(p_0). We then state and prove the main theorem.

Recall that if x_0 solves (1.1)(p_0), it is called a regular solution if $f_x(p_0,x_0)$ exists and if the linear generalized equation

$$y \in f(p_0,x_0) + f_x(p_0,x_0)(x-x_0) + \partial\psi_K(x) \tag{3.1}$$

defines a (single-valued) Lipschitzian function $x = x(y)$ from a neighborhood of the origin to a neighborhood of x_0 (see [2] and [3]; in [2] this property was called "strong regularity"). It turns out that we can take advantage of the structure of K near x_0 to simplify (3.1) and to make the property of regularity geometrically clearer.

First, since x_0 solves (1.1)(p_0) we know that $0 \in f(p_0,x_0) + \partial\psi_K(x_0)$, and thus the set $K_0 := \{x \in K \mid \langle f(p_0,x_0), x-x_0\rangle = 0\}$ is a face of K. Since K is polyhedral, for any h near $f(p_0,x_0)$ the set of points x satisfying $0 \in h + \partial\psi_K(x)$ and that satisfying $0 \in h + \partial\psi_{K_0}(x)$ are the same [4, Lemma 3.5]. But, again by polyhedrality, near x_0 the face K_0 coincides with $x_0 + T$, where T is the tangent cone to K_0 at x_0. Therefore, for h near $f(p_0,x_0)$ and v near the origin, we have $-h \in \partial\psi_K(x_0 + v)$ if and only if $-h \in \partial\psi_T(v)$. With these observations, we can use the following proposition to reduce the question of regularity to one of unique solvability for a simpler generalized equation.

PROPOSITION 3.1: Suppose x_0 solves (1.1)(p_0) and $f_x(p_0,x_0)$ exists. Then x_0 is a regular solution if and only if the generalized equation

$$y \in f_x(p_0,x_0)v + \partial\psi_T(v) \tag{3.2}$$

has a unique solution $v = v(y)$ for each $y \in R^n$. Further, if this is so then v is Lipschitzian, and for each y near 0 the solutions $x(y)$ and $v(y)$ of (3.1) and (3.2) respectively satisfy $x(y) = x_0 + v(y)$.

Proof: We first show that for y near 0 and x near x_0, and with $v = x-x_0$, the pair (y,x) satisfies (3.1) if and only if the pair (y,v) satisfies (3.2). If (y,v) is near $(0,0)$ and satisfies (3.2), then $v \in T$. However, by construction $f(p_0,x_0)$ is orthogonal to each element of T, and thus (3.2) is equivalent to

$$y \in f(p_0,x_0) + f_x(p_0,x_0)v + \partial\psi_T(v) \quad . \tag{3.3}$$

for x near x_0 we know $\partial\psi_{K_0}(x) = \partial\psi_T(x-x_0)$, so we can rewrite (3.3) with $v = x-x_0$ as

$$y \in f(p_0,x_0) + f_x(p_0,x_0)(x-x_0) + \partial\psi_{K_0}(x) \quad . \tag{3.4}$$

Further, since y is also near the origin the quantity $-y + f(p_0,x_0) + f_x(p_0,x_0)(x-x_0)$ is close to $f(p_0,x_0)$. But then by [4, Lemma 3.5] the relation (3.4) holds if and only if the same relation holds with K in place of K_0: i.e., if and only if we have (3.1).

Conversely, if (y,x) is near $(0,x_0)$ and satisfies (3.1), then we know $x \in K$; as y is near 0 and x is near x_0 we can reverse the previous argument to obtain (3.4) and then (3.3). This tells us that $v = x-x_0$ must belong to T, and we can then follow the same argument back to obtain (3.2).

Now suppose that (3.2) has a unique solution $v(y)$ for each $y \in \mathbb{R}^n$. The function $v(y)$ is polyhedral, hence everywhere locally upper Lipschitzian with some constant modulus [6, Prop. 1]. But by hypothesis v is single-valued, and it is easy to show that it is then actually Lipschitzian with the same modulus. It follows that, since $(0,0)$ satisfies (3.2), if y is any point near 0 then $v(y)$ is also small, and hence $x = x_0+v(y)$ is near x_0. By our previous argument the pair (x,y) satisfies (3.1). If x' is near x_0 and (y,x') satisfies (3.1), then again we apply the previous argument to conclude that $(y,x'-x_0)$ satisfies (3.2). By uniqueness of v, we then have $x'-x_0 = x-x_0$, so that x is locally unique as a solution of (3.1). But x inherits the Lipschitz modulus of v; hence x_0 is a regular solution of (1.1)(p_0).

Finally, suppose x_0 is a regular solution of (1.1)(p_0). Let $y \in \mathbb{R}^n$, and let $\alpha > 0$ be such that αy is close enough to 0 for the definition of regularity to yield a locally unique x near x_0 with $(\alpha y,x)$ satisfying (3.1). Since regularity requires that x be Lipschitzian, we can suppose α is so small that the pair $(\alpha y,x)$ is close enough to $(0,x_0)$ for our previous argument to apply. That argument shows that if we define v by $\alpha v = x-x_0$, then $(\alpha y,\alpha v)$ satisfies (3.2). However, since T is a cone the pair (y,v) then also satisfies (3.2). If v' is such that (y,v') satisfies (3.2), then for small positive β the pairs $(\beta y,\beta v)$ and $(\beta y,\beta v')$ both satisfy (3.2) and are close to $(0,0)$. Letting $x = x_0+\beta v$ and $x' - x_0+\beta v'$, we apply the previous argument to show that $(\beta y,x)$ and $(\beta y,x')$ each satisfy (3.1). But by the local uniqueness requirement of the regularity assumption, we have $x = x'$; thus $v = v'$ and the solution of (3.2) is globally unique. This proves Proposition 3.1.

We are now ready to prove the main result, which says that the solution of (3.2) for a special choice of y yields the B-derivative of $x(\cdot)$ at p_0. For completeness we include in the statement of the theorem some results from [2], establishing existence and Lipschitz continuity of the implicit function defined by (1.1)(p), but the result on B-differentiability is our main object here.

THEOREM 3.2: Let f be a Lipschitzian function from $\Pi \times \Omega \to \mathbb{R}^n$. Suppose that:

a. The partial F-derivative f_x exists on Ω and is continuous at (p_0,x_0), and

b. x_0 is a regular solution of (1.1)(p_0).

Then there are neighborhoods N_x of x_0 and N_p of p_0, and a Lipschitzian function $x : N_p \to N_x$, such that for each $p \in N_p$, $x(p)$ is the unique solution in N_x of (1.1)(p). Further, the B-derivative of x at p_0 is given by $Dx(p_0)(w) = v_L$, where $v_L = v_L(w)$ is the unique solution of the linear generalized equation

$$0 \in f_x(p_0,x_0)v_L + D_pf(p_0,x_0)(w) + \partial\psi_T(v_L) \quad (3.5)$$

Proof: The conclusion about $x(\cdot)$, except for the B-derivative, follows from [2, Th. 2.1, Cor. 2.2]. Since x is Lipschitzian on N_x, [5, Th. A.2] applies to show that it will be B-differentiable at p_0 with B-derivative $v_L(w)$ given by (3.5) provided we can show that $v_L(\cdot)$ is a single-valued function whose graph is a cone and that for w near 0, $x(p_0+w) = x_0+v_L(w)+o(w)$. The single-valuedness follows from our comments on (3.2), since (3.5) is nothing but (3.2) with $y = -D_pf(p_0,x_0)(w)$. Also, since $D_pf(p_0,x_0)(\cdot)$ is positively homogeneous and since T is a cone, the graph of $v_L(\cdot)$ is a cone. Therefore we have only to establish the approximation property.

Recall that for p near p_0 we have

$$0 \in f(p,x(p)) + \partial\psi_K(x(p)) \quad . \quad (3.6)$$

As $x(p)$ remains near x_0, by reasoning as in the proof of Proposition 3.1 we can show that (3.6) is equivalent to

$$0 \in f(p_0+w,\ x_0+v(w)) + \partial\psi_T(v(w)) \quad , \tag{3.7}$$

where $w = p-p_0$ and $v(w) = x(p_0+w) - x_0$. This in turn can be rewritten as

$$-f(p_0+w,x_0+v(w)) + f_x(p_0,x_0)v(w) \in f_x(p_0,x_0)v(w) + \partial\psi_T(v(w)) \quad . \tag{3.8}$$

We can also rewrite (3.5), recalling that $f(p_0,x_0)$ is orthogonal to each element of T, to obtain

$$-f(p_0,x_0) - D_pf(p_0,x_0)(w) \in f_x(p_0,x_0)v_L + \partial\psi_T(v_L) \quad . \tag{3.9}$$

Using the Lipschitz property of (3.2) and writing λ for the modulus, we find from (3.8) and (3.9) that

$$\begin{aligned}\|v(w)-v_L(w)\| &\leq \lambda\|f(p_0+w,x_0+v(w)) - f(p_0,x_0) \\ &\quad - f_x(p_0,x_0)v(w) - D_pf(p_0,x_0)(w)\| \\ &= o\left(\begin{bmatrix} v(w) \\ w \end{bmatrix}\right) \quad ,\end{aligned}$$

where we have used Proposition 2.1. However, since $x(\cdot)$ is Lipschitzian in p, $v(\cdot)$ is Lipschitzian in w. It follows that $\|v(w) - v_L(w)\| = o(w)$, which completes the proof of Theorem 3.2.

REFERENCES

[1] J. Kyparisis, "Sensitivity analysis framework for variational inequalities", preprint, 1985.

[2] S. M. Robinson, "Strongly regular generalized equations", Mathematics of Operations Research 5(1980) 43-62.

[3] S. M. Robinson, "Generalized equations", in: A. Bachem, M. Grötschel, and B. Korte, eds., Mathematical Programming: The State of the Art, Bonn 1982 (Springer-Verlag, Berlin 1983) pp. 346-367.

[4] S. M. Robinson, "Local structure of feasible sets in nonlinear programming, Part II: Nondegeneracy", Mathematical Programming Study 22(1984) 217-230.

[5] S. M. Robinson, "Local structure of feasible sets in nonlinear programming, Part III: Stability and sensitivity", accepted by Mathematical Programming Studies.

[6] S. M. Robinson, "Some continuity properties of polyhedral multifunctions", Mathematical Programming Study 14(1981) 206-214.

SMR/jvs

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 2854	2. GOVT ACCESSION NO. AD-A160 980	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Implicit B-Differentiability in Generalized Equations		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Stephen M. Robinson		8. CONTRACT OR GRANT NUMBER(s) DCR-8502202 DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53705		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 5 - Optimization and Large Scale Systems
11. CONTROLLING OFFICE NAME AND ADDRESS See Item 18 below		12. REPORT DATE August 1985
		13. NUMBER OF PAGES 9
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, DC 20550

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Equilibrium, generalized equations, implicit functions, Bouligand derivative

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

We have previously proved an implicit-function theorem for regular solutions of generalized equations. Here we show that when the underlying set for the generalized equation is polyhedral, as it is in many applications, then the implicit function has a Bouligand derivative defined by a formula generalizing that of the usual implicit-function theorem. This extends recent results on directional differentiability by Kyparisis and others.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

END

FILMED

12-85

DTIC